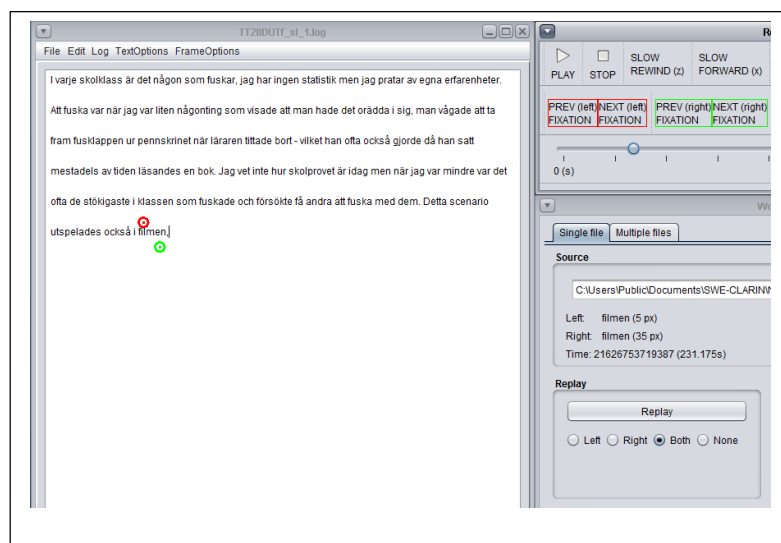


Humanistlaboratoriet vid Lunds universitet är en toppmodern forsknings- och utbildningsenhet knuten till humanist- och teologfakulteterna. Labbet har bl a utrustning för ögonrörelsemätning, artikulografi, motion capture, virtual reality och EEG. Labbet tillhandahåller även olika typer av teknologier och utbildningar för humaniora, däribland språkarkivering och träning i korpora- och datahantering.

Vid Humanistlaboratoriet utvecklar vi ScriptLog, ett verktyg för skrivforskning. Med ScriptLog är det möjligt att logga skrivprocessen medan man skriver. Denna kan sedan spelas upp och analyseras i efterhand. Vi har även möjlighet att koppla på ögonrörelsemätning för att se vilken del av texten som skribenten tittar på vid ett visst tillfälle. Figuren visar programmet vid en uppspelning av en text och ögonrörelsemätning där skribenten just fixerat på ordet 'filmen'.



ScriptLog utvecklas i samband med bl a Institutionen för svenska språket vid Göteborgs Universitet och Avdelningen för humanvetenskap vid Högskolan i Kristianstad. Det används i flera projekt inom språkvetenskap och logopedi samt även på grund- och gymnasieskolor.

Vid Humanistlaboratoriet pågår ett stort antal projekt som genererar data i många olika former. Humanistlaboratoriet tillhandahåller och underhåller en

korpusserver, <https://corpora.humlab.lu.se/>, där olika projekt kan lagra sina data. En viktig del i att beskriva innehållet och/eller strukturen för en viss datasamling är metadata. Kontinuerligt arbete pågår med att lägga in material från olika forskargrupper och -projekt samt att medverka till annotering och metadatauppmärkning av detta.

Metadata sker i modellen CMDI, som utvecklats inom CLARIN centralt. För att åskådliggöra materialet på servern finns även en korpusbrowser, som möjliggör sökning och visning av metadata för olika typer av material. Figuren visar en del av en korpus från språket Khmu, som talas i Laos, samt hur det annoteras med verktyget Elan.



Swe-Clarins första HS-workshop

Fredagen den 17 april hölls Swe-Clarins första workshop för humaniora och samhällsvetenskap (HS). Inbjudan var öppen för alla men riktad framförallt till forskare inom HS på Göteborgs universitet. Workshopen samlade omkring 50 deltagare och bestod av tre sessioner. Den första behandlade pågående (pilot-)projekt inom Swe-Clarin-konsortiet, den andra rörde material som finns eller är på gång inom Swe-Clarin och programmet avslutades med en diskussionssession.

Under första sessionen beskrev Jonas Lindström från Uppsala universitet under titeln "Att kika in i mörkret" nyttan av språkteknologi för projektet Gender and Work, där ett stort historiskt material undersöks för att hitta omnämningen av aktiviteter utförda av kvinnor och män. Projektet är ett samarbete mellan historiker och språkteknologer. Linn Sandberg från Universitetet i Oslo presenterade ett projekt om val- och opinionsforskning på Twitter-data i ett samarbete mellan statsvetenskap och Språkbanken på Göteborgs universitet samt datavetenskap på Chalmers, där fokus var att använda maskininlärning för att undersöka vilka ämnen som diskuteras på Twitter i samband med TV-sända Agendadebatter. Vidare berättade Leif-Jöran Olsson från Göteborgs universitet om Dramawebben och dess potential, dels som ett välannoterat material för humanistiska frågeställningar men även som tränings- och testmaterial för maskininlärning av yrken och roller. Slutligen presenterade Nina Tahmassebi från Göteborgs universitet tre projekt rörande relationsextraktion, svenska pseudosamordningar och retorikhistoria.

Under den andra sessionen diskuterades existerande samt påtänkt material. Lars Borin introducerade Språkbankens material och dess potential för frågeställningar inom HS. Jonas Engman berättade om Nordiska museets etnografiska frågelistor och lämnade under högtidliga former över dessa till Språkbanken. Anne Palmér och Annika Hillbom diskuterade en korpus av elevtexter under uppbyggnad vid Uppsala universitet (material ur nationella prov) och bad om feedback för utformningen för ett sådant material samt framtida frågeställningar. Slutligen informerade Lars Björk om ett projekt på Kungliga biblioteket för digitalisering av ett stort modernt bokmaterial.

Under diskussionssessionen berördes ämnen som material, samarbeten med biblioteken och frågeställningar. Bland annat efterfrågades ekonomiskt textmaterial, exempelvis årsrapporter, för samhällsvetenskaplig och ekonomisk forskning. Ett önskemål om en språkteknologisk terminologilista för HS-forskare som hjälpmedel för diskussioner med språkteknologer lades fram.

Workshopen avslutades med ett mingel. Många deltagare stannade kvar och fortsatte diskussionerna även efter de schemalagda sessionerna var över. Allt som allt var det en uppskattad workshop och vi ser fram emot nästa evenemang under hösten 2015.

Kalender

8-10 juni: Fonetik 2015, Lund,
<http://konferens.ht.lu.se/fonetik-2015/>
 9-10 juni: Nordic Clarin Network-workshop.
 5-6 oktober: Nordic Clarin Network-workshop i anslutning till Språkbankens höstworkshop.
 11 november: SNDs höstworkshop
 19-20 november: Swe-Clarins stormöte i Stockholm, Swe-Clarins-workshop i anslutning på fredag eftermiddag. Alla är välkomna.

Partners

Swe-Clarins har nio partners från Lund, Göteborg, Linköping, Stockholm och Uppsala, både på universitet och myndigheter. En lista och beskrivning av alla partners hittar ni här:
<http://sweclarin.se/swe/centrum>

Nyheter

Vi kommer inte att fortsätta spamma. Om ni vill få vidare information om Swe-Clarins anmäl er till e-postlistan här:
http://lists.sweclarin.se/mailman/listinfo/news_lists_sweclarin.se

